The revised manuscript is much improved. Well done by the authors. While it could be on track for publication, there are a few important considerations and perhaps one methodological error that must first be addressed. Please see my comments below, which are intended to be positive and constructive. I would be happy to review a revised version of the manuscript and I again commend the authors for the great progress they have made with this manuscript.

Major comments

1. What were the units for discharge when you modeled constituent fluxes with LOADEST? From your SI dataset, it seems that you used $m^3\ s^{-1}$? As detailed in Runkel et al. (2004), LOADEST takes discharge in units of $ft^3\ s^{-1}$. If you used $m^3\ s^{-1}$, then your constituent modeling must be updated and associated values throughout the text corrected (e.g., Sec. 3.1). Of course, this would be relatively easy to do! This being said, I agree with reporting discharge in metric units ($m^3\ s^{-1}$) in the main text (e.g., L210), as is commonly done.

   Runkel, R. L., Crawford, C. G., & Cohn, T. A. (2004). *Load Estimator (LOADEST): A FORTRAN program for estimating constituent loads in streams and rivers* (No. 4-A5). https://pubs.usgs.gov/tm/2005/tm4A5/pdf/508final.pdf

2. Following on Dr. Park's comment about headwater streams: Headwaters are generally the smallest surficial fluvial component within a stream network. I read Meyer and Wallace (2001) and could not find where they explicitly state that headwaters include 1st to 3rd order streams, as your citation in L34 implies… but please feel free to prove me wrong about this :) ! The lowest stream order is not necessarily a headwater stream, because definitions of stream order are relative and can vary depending on the resolution of the data used to determine stream order. For instance, stream order derived from geospatial data like a DEM (often relatively coarse, e.g., 90 m from Lin et al., 2019 WRR) may be quite different than stream order determined by an expert in the field or using high-resolution satellite imagery to visually map streams. From Google Earth imagery and your Figure 1, sampling points SLH2–5 do not appear to be headwater streams, and I am not convinced that SLH0–1 are, either. All this to say, I am not convinced that the streams in your study are truly headwater streams and I would not use GIS-derived estimates of stream order (or the citation to Meyer and Wallace 2001) to try to assert this. I think it would be appropriate to place less emphasis on interpreting your results as direct measurements of headwater streams. I think this does not require a major rewrite of your paper and can be addressed without too much trouble by modifying the very nice text that you have in the Introduction and elsewhere. For instance, you could reason that your sampling points capture the integrated hydrochemical signal across headwater catchments. This would be compelling and more accurate. Clarifying this, and omitting 'headwater' where appropriate, would more accurately represent your study sites without comprising the overall relevance of your research, or its significance for understanding hydrochemical signals within headwaters of the QTP.

Minor comments

1. Also related to LOADEST: I see now (from the text in Sec. 2.5 and Table S1) that you consider all possible LOADEST regression equations when modeling C flux. i.e., It appears that you selected Model 0 and let LOADEST automatically choose the most parsimonious model

(lowest AIC) from models 1–9. However, model #s 3, 5, 7–9 include covariates for long-term change, which are arguably more appropriate to use when you have measurements from across multiple years (even five years of observations is not considered long enough by some studies, e.g., Zolkos et al., 2020). Because your study period is not multi-annual, and you are not investigating long-term trends, it would be more appropriate to estimate fluxes using the best model from 1, 2, 4, or 6.

Zolkos, S., Krabbenhoft, D. P., Suslova, A., Tank, S. E., McClelland, J. W., Spencer, R. G., ... & Holmes, R. M. (2020). Mercury Export from Arctic Great Rivers. *Environmental Science & Technology*, *54*(7), 4140-4148.

Additional comments

L14: “dynamics” is vague. Avoiding it generally improves clarity. If there is more to the story than just C ‘transport’ / fluxes, consider clarifying how ‘dynamics’ relates to it.

L17: What do you mean by “in-depth”? Please clarify.

L22: “… thawed frozen…”. Are the soils thawed or frozen? I see what you are trying to say here. It would be clearer to just say ‘thawed’.

L30: “dynamics”- see comment for L14.

L30: What exactly do you mean by “cascading effects”? i.e., that thaw effects in headwaters might propagate downstream, across watershed scales? (especially for dissolved inorganic carbon, in your region) If so, you might consider recent work from the western Canadian Arctic exploring this topic: https://bg.copernicus.org/articles/17/5163/2020/bg-17-5163-2020.html, but also recognize that thaw effects on fluvial C cycling (especially DIC) across watershed scales vary across permafrost regions. For Siberia, see:
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2017JG004311
Perhaps worth considering in the Discussion, and how your work adds to this understanding.

L63-4: Why is it important that there are different C sources? It would be interesting and helpful to clarify the biogeochemical relevance of the C sources, and doing so could help to justify the rationale for your Ad/Al ratio measurements. For instance, do the sources have varying degrees of recalcitrance?

L89: The value for mean annual daily discharge is interesting. It would also be interesting to know total annual discharge (e.g., $km^3\ y^{-1}$), from your daily Q measurements.

L90-1: Please see my Major comment #2.

L95-6: No carbonate lithologies? Why are your DIC concentrations so high? This should be discussed a bit more. My recollection is that some of the recent work by Song et al. discusses carbonate lithologies on the QTP.

L202: Please specify: What post-hoc test did you use?

L239: Would be good to consistently include valencies for calcium and magnesium throughout, as you do here (e.g., L156).

L307: Pardon the editorial comment, but it would be more appropriate for dissolved ‘bicarbonates’ and ‘carbonates’ to be singular, not plural :)

L311-12: Please see my comment for L95-6. Are there carbonate lithologies? Consistency in the message would help.

L313: Technically, high Ca^{2+} and Mg^{2+} concentrations and a strong association with DIC reflect greater availability of carbonate-bearing lithologies for chemical weathering, not necessarily high rates of weathering. So, I am not sure that I would say “proved” here. Perhaps say “reflected” or “indicated by”.

L318-320: Rainfall should enhance carbonate weathering, no? (e.g., Zeng et al., 2019) Perhaps clarify and elaborate on this.

Zeng, S., Liu, Z., & Kaufmann, G. (2019). Sensitivity of the global carbonate weathering carbon-sink flux to climate and land-use changes. *Nature Communications*, *10*(1), 1-10.

L352: “headstream”- i.e., “headwater stream”?

L398: You could simply and clarify by saying “DIC” instead of “dissolved carbon (especially DIC)”.

L401: Suggest deleting “frozen” (related to my comment for L22).

L402: “levels” = “concentrations”?

L402: Apparently higher sensitivity of DIC concentration than DOC concentration, but also very interesting that DOC source appears to change. Worth mentioning here.

Figure 2: Are the *y*-axis units metric ‘tonnes’ per day, rather than imperial ‘tons’? For easier comparison with other figures and values, it would help state the *y*-axis in metric units (e.g. Megagrams, Mg, or in grams if desired) and re-scale as needed.